Local indicators of spatio-temporal association: LISTA functions

Francisco J. Rodríguez-Cortés[1,*], Mohammad Ghorbani[2] and Jorge Mateu[1]

[1] *Department of Mathematics, Universitat Jaume I, Castellón, Spain; cortesf@uji.es, mateu@uji.es*
[2] *Department of Mathematical Sciences, Aalborg University, Denmark; ghorbani@math.aau.dk*
[*] *Corresponding author*

Abstract. *We consider here the problem of detecting local clusters in spatio-temporal point patterns. These tools are based on local second-order characteristics of spatio-temporal point processes. We extend the notion of spatial dependence to spatio-temporal structures defining the LISTA functions derived from second-order spatio-temporal product densities. We derive some theoretical properties, propose unbiased edge-corrected estimators and use these new functions to analyse and detect clustering in the spatio-temporal evolution of a disease.*

Keywords. *Clustering; Local indicators of spatio-temporal association; Second-order spatio-temporal product densities; Spatio-temporal K-function.*

---

1 Introduction

LISA functions are built from local second-order characteristics of spatial point processes through considering individual contributions from second-order product density functions (Cressie and Collins, 2001a,b). Rodríguez-Cortés *et al.* (2012) and Mateu and Rodríguez-Cortés (2014) introduced the first approaches for extending the concept of LISA to the spatio-temporal context. We here define a new version of these functions based on the most recent work by Gabriel and Diggle (2009), Møller and Gorbani (2012), Ghorbani (2013), Gabriel (2013) and Rodríguez-Cortés *et al.* (2014), and present edge-corrected estimators developing their first-order theoretical moment. An application to detect spatio-temporal clusters in public health problems is also considered.

Definitions and notations used throughout this paper are introduced by Møller and Gorbani (2012) and Rodríguez-Cortés *et al.* (2014). We consider a spatio-temporal point process with no multiple points as a random countable subset X of $\mathbb{R}^2 \times \mathbb{R}$, where a point $(\mathbf{u}, s) \in X$ corresponds to an event at $\mathbf{u} \in \mathbb{R}^2$ occurring at time $s \in \mathbb{R}$. In practice, we observe n events $\{(\mathbf{u}_i, s_i)\}$ of X within a bounded spatio-temporal region $W \times T \subset \mathbb{R}^2 \times \mathbb{R}$, with area $|W| > 0$, and length $|T| > 0$. Let N_{lfs} and N_{lft} be the spaces of locally finite subsets of $\mathbb{R}^2$ and $\mathbb{R}$ equipped with σ-algebras $\mathcal{N}_{\text{lfs}}$ and $\mathcal{N}_{\text{lft}}$ respectively, see Møller and Waagepetersen (2004). In the sequel, $N(A)$ denotes the number of the events of the process falling in a bounded region $A \subset W \times T$. For a given event $(\mathbf{u},s)$, the events that are *close* to $(\mathbf{u},s)$ in both space and time, for each spatial distance r, and time lag t, are given by the corresponding spatio-temporal cylindrical neighborhood of the event $(\mathbf{u},s)$, which can be expressed by the cartesian product as $b((\mathbf{u},s),r,t) = \{(\mathbf{v},l) : \|\mathbf{u}-\mathbf{v}\| \leq r, |s-l| \leq t\}$, with $(\mathbf{u},s),(\mathbf{v},l) \in \mathbb{R}^2 \times \mathbb{R}$, and where $\|\cdot\|$ denotes the Euclidean distance in $\mathbb{R}^2$ and $|\cdot|$ denotes the usual distance in $\mathbb{R}$. Note that $b((\mathbf{u},s),r,t)$ is a cylinder with center $(\mathbf{u},s)$, radius r and height $2t$.

Assume that $\rho(\mathbf{u},s)$ is the spatio-temporal intensity, and $\rho^{(2)}((\mathbf{u},s),(\mathbf{v},l))$ the second-order product density function. A process for which $\rho(\mathbf{u},s) = \rho$ for all $(\mathbf{u},s) \in X$ is called homogeneous of first-order. Further, if $\rho^{(2)}((\mathbf{u},s),(\mathbf{v},l)) = \rho^{(2)}(\mathbf{u}-\mathbf{v},s-l)$, the process is called second-order or weak stationary (Ghorbani, 2013). We assume that the point process X is orderly, roughly meaning that coincident points cannot occur. We assume first- and second-order spatio-temporal separability hypothesis, i.e.,

$$\rho(\mathbf{u},s) = \bar{\rho}_1(\mathbf{u})\bar{\rho}_2(s), \quad (\mathbf{u},s) \in \mathbb{R}^2 \times \mathbb{R}, \tag{1}$$

and

$$\rho^{(2)}((\mathbf{u},s),(\mathbf{v},l)) = \bar{\rho}_1^{(2)}(\mathbf{u},\mathbf{v})\bar{\rho}_2^{(2)}(s,l), \quad (\mathbf{u},s),(\mathbf{v},l) \in \mathbb{R}^2 \times \mathbb{R} \tag{2}$$

where $\bar{\rho}_1, \bar{\rho}_2, \bar{\rho}_1^{(2)}, \bar{\rho}_2^{(2)}$ are non-negative functions. For more details see Møller and Gorbani (2012) and Rodríguez-Cortés *et al.* (2014). Considering the hypothesis of first-order spatio-temporal separability in (1), we have that $\rho(\mathbf{u},s) = (\rho_{\text{space}}(\mathbf{u})\rho_{\text{time}}(s))/\int \rho(\mathbf{u},s)\,\mathrm{d}(\mathbf{u},s)$. For a stationary point process X, ρ, ρ_{space} and ρ_{time} are all constant. For non-parametric estimation of ρ_{space}, ρ_{time} and $\rho(\mathbf{u},s)$, see Ghorbani (2013).

Throughout this paper we assume that X is second-order intensity-reweighted stationary (SOIRS), i.e. $\rho^{(2)}((\mathbf{u},s),(\mathbf{v},l)) = \rho^{(2)}(\mathbf{u}-\mathbf{v},s-l)$, with $(\mathbf{u},s),(\mathbf{v},l) \in \mathbb{R}^2 \times \mathbb{R}$ (Baddeley *et al.* (2000), Gabriel and Diggle (2009)). Further, if the process is isotropic, then $\rho^{(2)}(\mathbf{u}-\mathbf{v},s-l) = \rho_0^{(2)}(\|\mathbf{u}-\mathbf{v}\|,|s-l|)$ for some non-negative function $\rho_0^{(2)}(\cdot)$. Just as in the spatio-temporal first-order case, considering the hypothesis of second-order spatio-temporal separability in (2), we have that

$$\rho^{(2)}((\mathbf{u},s),(\mathbf{v},l)) = \frac{\rho^{(2)}_{\text{space}}(\mathbf{u}-\mathbf{v})\rho^{(2)}_{\text{time}}(s-l)}{\int\int \rho^{(2)}(\mathbf{u}-\mathbf{v},s-l)\,\mathrm{d}(\mathbf{u},s)\,\mathrm{d}(\mathbf{v},l)}. \tag{3}$$

For an unbiased estimator of (3) and its properties of the second-order spatio-temporal product density function, see Rodríguez-Cortés *et al.* (2014).

For a SOIRS and isotropic spatio-temporal point process X, Gabriel and Diggle (2009) extended the Ripley's K-function to the spatio-temporal inhomogeneous K-function. For a Poisson process, $K(r,t) = 2\pi r^2 t$. For an unbiased estimator of the K-function, see Gabriel (2013). Both in the stationary and isotropic case and, under SOIRS and isotropic case, the second-order spatio-temporal product density function is proportional to the derivative of $K(r,t)$ with respect to r and t, i.e. in the planar case,

$$\rho^{(2)}(r,t) = \frac{\rho(\mathbf{u},s)\rho(\mathbf{v},t)}{4\pi r}\frac{\partial^2}{\partial r \partial t}K(r,t).$$

For a stationary point process X, $\rho(\mathbf{u},s) = \rho$ and $\rho^{(2)}(r,t) = (\rho^2/4\pi r)\partial^2 K(r,t)/\partial r \partial t$, where $\rho^2 K(r,t)$ is the expected number of ordered pairs of distinct points per unit volume of the observation window with pairwise distance and time lag less than r and t (Rodríguez-Cortés *et al.*, 2014).

Under the stationarity case and ignoring edge-effects, a global naive non-parametric kernel estimator for $\rho^{(2)}(r,t)$ in (3) is given by

$$\widehat{\rho^{(2)}}_{\varepsilon,\delta}(r,t) = \frac{1}{4\pi r|B|}\sum_{i=1}^{n}\sum_{j\neq i}\kappa_{\varepsilon,\delta}(\|\mathbf{u}_i-\mathbf{u}_j\|-r,|s_i-s_j|-t), \tag{4}$$

with $r>\varepsilon>0, t>\delta>0$ and $B=W\times T$. We assume that the kernel function κ has the multiplicative form $\kappa_{\varepsilon\delta}(\|\mathbf{u}_i-\mathbf{u}_j\|-r,|s_i-s_j|-t)=\kappa_{1\varepsilon}(\|\mathbf{u}_i-\mathbf{u}_j\|-r)\,\kappa_{2\delta}(|s_i-s_j|-t)$, where $\kappa_{1\varepsilon}$ and $\kappa_{2\delta}$ are respectively kernel functions with bandwidths ε and δ. Both the K-function and the product density function provide a global measure of the covariance structure by summing over the contributions from each event observed in the process.

2 LISTA functions

For a stationary and isotropic spatio-temporal point process X, we can define a local version of the K-function as $\{\rho K(r,t)\}^i = \mathbb{E}[N(b((\mathbf{u}_i,s_i),r,t)\setminus\{(\mathbf{u}_i,s_i)\})|(\mathbf{u}_i,s_i)\in X]$, with $r>0, t>0$, where the expectation is conditional on observing $(\mathbf{u}_i,s_i)\in X$ and calculated with respect to the reduced Palm measure. This can be interpreted as the expected number of extra events from $(\mathbf{u}_i,s_i)$ with pairwise distance and time lag less than r and t respectively.

The local indicator of spatio-temporal association (LISTA) is a local function which considers individual points. We denote the localised version of the second-order product density by $\rho^{(2)i}$. A kernel estimate of $\rho^{(2)i}$ is given by

$$\widehat{\rho^{(2)i}}_{\varepsilon,\delta}(r,t) = \frac{n-1}{4\pi|B|r}\sum_{i\neq j}\kappa_{1\varepsilon}(\|\mathbf{u}_i-\mathbf{u}_j\|-r)\kappa_{2\delta}(|s_i-s_j|-t), \tag{5}$$

with $r>\varepsilon>0, t>\delta>0$. For fixed r and t it holds that

$$\widehat{\rho^{(2)}}_{\varepsilon,\delta}(r,t) = \frac{1}{n-1}\sum_{i=1}^{n}\widehat{\rho^{(2)i}}_{\varepsilon,\delta}(r,t).$$

Following Cressie and Colins (2001a,b), we have that for a homogeneous Poisson process

$$\begin{aligned}\mathbb{E}^{!}\left[\widehat{\rho^{(2)i}}_{\varepsilon,\delta}(r,t)\right] &= \frac{1}{4\pi|B|r}\mathbb{E}^{!}\left[(N(B)-1)\sum_{j\neq i}\kappa_{\varepsilon,\delta}(\|\mathbf{u}_i-\mathbf{u}_j\|-r,|s_i-s_j|-t)\right]\\ &= \frac{\rho^2+\frac{\rho}{|B|}}{4\pi r}\int_{r-\varepsilon}^{r+\varepsilon}\kappa_{1\varepsilon}(u-r)|\partial b(\mathbf{u}_i,u)\cap W|_1\mathrm{d}u\int_{t-\delta}^{t+\delta}\kappa_{2\delta}(v-t)|\partial b(s_i,v)\cap T|_0\mathrm{d}v,\end{aligned}$$

here $|\cdot|_0$ and $|\cdot|_1$ are the zero- and one-dimensional Hausdorff measures respectively in $\mathbb{R}$ and $\mathbb{R}^2$.

The corresponding edge-corrected second-order product density LISTA function is

$$\widehat{\rho^{(2)i}}_{\varepsilon,\delta}(r,t) = \frac{n-1}{4\pi|B|r}\sum_{i\neq j}\frac{\kappa_{1\varepsilon}(\|\mathbf{u}_i-\mathbf{u}_j\|-r)\kappa_{2\delta}(|s_i-s_j|-t)}{w^2(\mathbf{u}_i,\mathbf{u}_j)w^1(s_i,s_j)}, \tag{6}$$

with $r > \varepsilon > 0$, $t > \delta > 0$, for $(\mathbf{u}_i, s_i) \in W \times T$ and $i = 1, \dots, n$. Here, $w^2(\mathbf{u}_i, \mathbf{u}_j)$ is the Ripley's isotropic edge-correction factor, and $w^1(s_i, s_j)$ is the temporal edge-correction factor. And the expected value for a homogeneous process is

$$\mathbb{E}^! \left[\widehat{\rho^{(2)i}}_{\varepsilon,\delta}(r,t) \right] = \frac{\rho^2 + \frac{\rho}{|B|}}{4\pi r} \int_{r-\varepsilon}^{r+\varepsilon} 2\pi u \kappa_{1\varepsilon}(u-r) \mathrm{d}u \int_{t-\delta}^{t+\delta} 2\kappa_{2\delta}(v-t) \mathrm{d}v = \rho^2 + \frac{\rho}{|B|},$$

Acknowledgments. Francisco J. Rodríguez-Cortés' research was supported by grant P1-1B2012-52. Mohammad Ghorbani's research was supported by the Center for Stochastic Geometry and Advanced Bioimaging, funded by a grant from the Villum Foundation. Jorge Mateu's research was supported by grant MTM2010-14961 from Ministery of Education.

References

[1] Anselin, L. (1995), Local Indicators of Spatial Association–LISA. *Geographical Analysis*, **27**: 93–115.

[2] Baddeley, A., Møller, J. and Waagepetersen, R. (2000). Non- and semi-parametric estimation of interaction in inhomogeneous point patterns, *Statistica Neerlandica* **54**: 329–350.

[3] Cressie, N.A.C. and Collins, L.B (2001a). Analysis of spatial point patterns using bundles of product density LISA functions, *Journal of Agricultural, Biological, and Environmental Statistics*, **6**: 118–135.

[4] Cressie, N.A.C. and Collins, L.B. (2001b), *Patterns in spatial point locations: Local indicators of spatial association in a minefield with clutter*. Naval Research Logistics, **48**: 333–347.

[5] Gabriel, E. (2013). Estimating second-order characteristics of inhomogeneous spatio-temporal point processes, *Methodology and Computing in Applied Probability* **16(1)**: 1–21.

[6] Gabriel, E. and Diggle, P.J. (2009). Second-order analysis of inhomogeneous spatio-temporal point process data, Statistica Neerlandica **63**: 43–51.

[7] Ghorbani, M. (2013). Testing the weak stationarity of a spatio-temporal point process, *Stochastic Environonmental Research and Risk Assessesment* **27**: 517–524.

[8] Mateu, J. and Rodríguez-Cortés, F.J. (2014). Local clustering in spatio-temporal point patterns, in E. Pardo-Igúzquiza, C. Guardiola-Albert, J. Heredia, L. Moreno-Merino, J. J. Durán and J. A. Vargas-Guzmán (eds), *Mathematics of Planet Earth, Lecture Notes in Earth System Sciences*, Springer Berlin Heidelberg, pp. 171–174. URL: http://dx.doi.org/10.1007/978-3-642-32408-6 40

[9] Møller, J. and Ghorbani, M. (2012). Aspects of second-order analysis of structured inhomogeneous spatio-temporal point processes, *Statistica Neerlandica* **66**: 472–491.

[10] Møller, J. and Waagepetersen, R.P. (2004). *Statistical Inference and Simulation for Spatial Point Processes*, Chapman and Hall/CRC, Boca Raton.

[11] Rodríguez-Cortés, F., Mateu, J. and Lorenzo, G. (2012). Spatio-temporal analogues to LISA functions. In Proceedings of the Sixth International Workshop on Spatio-Temporal Modelling (METMA6). Menezes, R. *et al.* (Eds.). ISBN: 978-989-97939-0-3.

[12] Rodríguez-Cortés, F.J., Ghorbani, M., Mateu, J., and Stoyan, D. (2014). On the expected value and variance for an estimator of the spatio-temporal product density function. Department of Mathematical Sciences, Aalborg University.